Final Report:
AFRL/Cornell Information Assurance Institute

AFOSR Grant F49620-00-1-209

15 March 2000 – 31 August 2002

Fred B. Schneider (Director)
Department of Computer Science
Cornell University
Ithaca, New York 14853
(607) 255-9221 (phone)
(607) 255-4428 (fax)
fbs@cs.cornell.edu

Abstract

The AFRL/Cornell Information Assurance Institute supports a broad spectrum of research aimed at developing a science and technology base to enhance information assurance and networked information systems trustworthiness—system and network security, reliability, and assurance. The institute also fosters closer collaborations between Cornell and AFRL researchers, as well as facilitating technology transfer and exposing Cornell researchers to problems facing the Air Force.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	13 Nov 2002	Final 15 Mar 2000 -- 31 Aug 2002

4. TITLE AND SUBTITLE
AFRL/Cornell Information Assurance Institute

5. FUNDING NUMBERS
G F49620-00-1-0209

6. AUTHOR(S)

Schneider, Fred B.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Cornell University
4130 Upson Hall
Ithaca, NY 14851

8. PERFORMING ORGANIZATION REPORT NUMBER
38094

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NM
4015 Wilson Boulevard, Room 713
Arlington, VA 22203-1954

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; distribution is Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

The AFRL/Cornell Information Assurance Institute supports a broad spectrum of research aimed at developing a science and technology base to enhance information assurance and networked information systems trustworthiness---system and network security, reliability, and assurance.
The institute also fosters closer collaborations between Cornell and AFRL researchers, as well as facilitating technology transfer and exposing Cornell researchers to problems facing the Air Force.

14. SUBJECT TERMS
System and network security, reliability, and assurance.

15. NUMBER OF PAGES
24

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

1 Introduction

The AFRL/Cornell Information Assurance Institute (IAI) was established at Cornell by an initial grant from AFOSR in March 2000. IAI was created as a prototype for a new mode of funding. And, after 2 years, the results of this experiment confirm that this new funding mode–granting research funds to a University center having close geographic and intellectual proximity to, but loose affiliation with, an AFRL laboratory—has enormous leverage:

- IAI funding has enabled some absolutely first-rate Computer Science research to be performed at Cornell.
- IAI funding has facilitated interactions between Cornell researchers and AFRL staff, with research at Cornell now having clear relevance to the research needs of the Air Force.

Specific research accomplishments supported under the auspices of IAI are summarized below (§2); details can be found in the publications listed at the end of this report (§4). Technology transitions and DoD interactions are also discussed (§3). Figure 1 lists those researchers at Cornell (along with their specializations) whose work has been supported, in part, by IAI.

2 Summary of Research Accomplishments

Scalable Fault-tolerant Systems (Birman, van Renesse). This effort has focused on scalability of the publish-subscribe paradigm and has interacted extensively with Rome/AFRL researchers to understand specific issues arising from application of the publish-subscribe paradigm within the JBI effort and within other related military systems. The most significant accomplishments include the development of (i) the Astrolabe scalable monitoring and management framework and (ii) the Astrocast publish-subscribe structure based on a novel Bimodal Multicast. Together, these bring a new and remarkably flexible way of implementing publish-subscribe services with good scaling properties, stability under stress, and a high quality of security.

Program Refinement Logic (Constable, Kreitz). Program Refinement Logic (PRL) is a logical programming environment that provides substantial automation in the design, coding, verification, and evolution of large software systems. It is based on the latest version of Nuprl proof development system. A first prototype of a formal digital library of algorithmic knowledge (FDL) has been completed. FDL provides an infrastructure for

Kenneth Birman:	Distributed computing, fault-tolerant network systems, distributed systems security, large-scale network applications.
Robert L. Constable:	Applied logic, automated reasoning, software assurance.
Alan Demers:	Database systems, database replication, and algorithms.
Johannes Gehrke:	Database systems and data mining.
Joseph Y. Halpern:	Reasoning about knowledge and uncertainty, distributed computing, security.
Dexter Kozen:	Proof carrying code, program logics, and semantics.
Christoph Kreitz:	Applied logic, automated reasoning, software assurance.
J. Gregory Morrisett:	Programming languages, compilers, distributed systems, language-based security.
Andrew Myers:	Programming languages, security, mobile code.
Robbert Van Renesse:	Distributed computing, fault-tolerant network systems, distributed systems security, large-scale network applications.
Fred. B. Schneider:	Distributed systems security and fault-tolerance, mobile code, concurrent programming.
Emin Gun Sirer:	Secure distributed systems, extensible operating systems, language-based security, automated testing.
Jayavel Shanmugasundaram:	Internet data management, database systems, and query-processing in emerging system architectures.

Figure 1: IAI Staff and Research Interests

verifying and synthesizing software systems by supporting the creation of certified algorithmic knowledge, the cooperation of multiple theorem proving systems, and flexible yet controlled access to the archived knowledge. Users may contribute library contents using the Nuprl, MetaPRL, JProver, and PVS theorem provers.

FDL has been used to support:

- code transformations that improve performance and enable protocols to be made adaptive while preserving functionality in connection with a self-adaptive task allocation manager to control processing of real-time media over a network through coordinated local schedules,
- the creation of formal courseware, and
- the translation of formal proofs into natural language.

Databases and Data Mining (Gehrke). The Cougar Project has produced database technology to support distributed wireless sensor networks with millions of nodes. Here, novel distributed query processing strategies for long-running queries permit in-network aggregation and can trade communication for local computation, increasing the lifetime of the network by up to an order of magnitude. A first version of the system was demonstrated at 29 Palms in California (Fall 2001) and the ACM Sigmod Conference (2002).

The Himalaya Project has created some of the world's fastest data mining algorithms for mining long itemsets, classification tree construction, and also regression-tree construction and sequence mining. Other work focused on pushing user-defined constraints (such as defined by intrusion-detection systems) deep into the mining algorithm, in order to improve performance by orders of magnitude versus simple a-priori model construction and a-posteriori model pruning via constraints. This project also investigated privacy-preserving data mining algorithms, where datasets can be shared publicly without compromising values of individual records while at the same time ensuring that accurate statistical summary information can still be recovered.

Formalizing Security (Halpern). The use of modal logic has led to a new formalization of secrecy. This formalization is consistent with Sutherland's notion of nondeducibility, subsumes both separability generalized non-interference and nondeducibility on strategies, and is able to handle

probabilistic secrecy, resource-bounded reasoning, as well as downgrading of information.

In addition, a new first-order logic was developed for reasoning about security policies. The formalism is a fragment of first-order logic that can both express many policies of interest and is tractable. Based on the logic, a prototype reasoning engine has been designed. Its user interface is intended for non-logicians, allowing them to enter policies, facts about principals, and then to ask questions about the policies.

Avoiding Malicious Boot Firmware (Kozen). In collaboration with Architecture Technologies Corporation (Ithaca, NY) and CodeGen Inc. (Palo Alto, CA) a prototype certifying compiler and verifier for detecting malicious boot firmware has been developed. Boot firmware modules are automatically verified against a standard security policy, as they are loaded. Among other things, the security policy being enforced asserts that drivers must access other devices only through a strict interface and may not access memory or bus addresses not allocated to them. Efficient Code Certification, along with inexpensive static checks on the compiled code, suffice to guarantee dynamic properties of the program at run time. The prototype is compliant with the now widely used IEEE 1275 Open Firmware standard for boot firmware. Sample device drivers written in Java for a block-oriented storage device and a PCI bus have been successfully compiled.

Cyclone Compiler (Morrisett). Cyclone is type-safe programming language that can be roughly characterized as a "superset of a subset of C." The type system of Cyclone accepts many C functions without change and uses the same data representations and calling conventions as C for a given type constructor. It also rejects many C programs to ensure safety. For instance, it rejects programs that perform (potentially) unsafe casts, that use unions of incompatible types, that (might) fail to initialize a location before using it, that use certain forms of pointer arithmetic, or that attempt to do certain forms of memory management. All of the analyses used by Cyclone are local (i.e., intra-procedural) to enable scalability and separate compilation. The analyses are carefully constructed to avoid unsoundness in the presence of threads.

Experimental validations of Cyclone show great promise. For systems applications, such as a simple web server, Cyclone introduces virtually no overhead at all. This is not surprising, as these applications tend to be I/O-bound. For scientific applications, a much larger overhead is seen (around

5x for a naive port, and 3x with an experienced programmer). Some of that overhead is due to bounds and null pointer checks on array access, which can be eliminated using intra-procedural analysis; other overhead arises from the use of "fat pointers" and the fact that GCC does not always optimize struct manipulation.

Secure Program Partitioning (Myers). The Jif/split prototypes bring a new means to ensure that data confidentiality and integrity are preserved in distributed systems in spite of untrusted hosts and mutually distrusting principals. This problem is particularly relevant to information systems used by mutually distrusting organizations, such as the dynamic coalitions that arise in military settings. With Jif/split, programs are automatically partitioned into communicating subprograms that run on the available, partially trusted hosts; to protect data integrity, information and code are also replicated across the available hosts. If any host is subverted, then only principals that have explicitly stated trust in that host need fear a violation of confidentiality.

Inlined Reference Monitors (Schneider, Morrisett). In-lined reference monitors are a new approach to implementing traditional reference monitors whereby a desired end-to-end security policy is formulated using a high-level declarative policy language and then a rewriting tool is used to automatically rewrite untrusted code into code that respects the policy. The rewriting tool works by inserting extra state and dynamic checks into the untrusted code so that the code becomes self-monitoring.

Having developed prototypes for Intel X86 and Java JVM, the central question is one of practicality. To this end:

- A set of kernel modifications was developed to support a prototype IRM rewriter in Microsoft's Windows operating system.
- A prototype MSIL (Microsoft Intermediate language) IRM realization has been developed. It implements an aspect-oriented programming metaphor for MSIL assembly language (rather than for a high-level language).

Internet Data Management and Retrieval(Shanmugasundaram). The QUARK project aims to integrate the database and information retrieval worlds by building a next-generation database system for handling both structured and unstructured data. This has required the development

of new techniques for storing and querying semi-structured data (containing a mix of structured and unstructured data) by using structured relational database systems. Techniques have also been developed for evaluating exploratory ranked keyword search queries over semi-structured data.

The PEPPER project has two main goals:

- to build an efficient information dissemination (or publish-subscribe) system for large-scale distributed systems and
- to develop a query processing layer for peer-to-peer networks.

The first goal required the development of a new, scalable index structure, called RPH-trees, for indexing user preferences so that only the relevant users are notified when new information becomes available. An interesting feature of RPH-trees is that they dynamically adapt to the information workload. For example, if there is a sudden burst of information about vehicle movement in Northern Afghanistan, the RPH-tree dynamically and automatically adjusts itself so that this information is processed efficiently and without delay. For the second goal, a fault-tolerant and scalable peer-to-peer index structure called P-trees has been developed. P-trees can support range queries in addition to equality queries.

MagnetOS (Sirer). MagnetOS is a new distributed operating system for ad hoc networks. It extends the effective lifetime of an ad hoc or sensor network through dynamic object migration, providing a single system image of a unified Java virtual machine across the nodes comprising an ad hoc network. By automatically and transparently partitioning applications into components and dynamically placing these components on nodes within the ad hoc network, MagnetOS reduces energy consumption, avoids hotspots, and increases system longevity—system longevity is increased by a factor of four to five, in fact.

Developing MagnetOS required solving two significant problems in ad hoc networks:

- **Multipath route selection**. Most routing algorithms—including those that are used in the core of the Internet—use single-path routing and thus are slow to respond to failures and frequently suffer path failures. Consequently, a new, efficient algorithm for constructing highly-reliable path sets has been developed.
- **Hybrid Routing Framework.** Traditional routing algorithms either proactively disseminate route updates or defer route discovery

until needed by a client. Choosing between the two regimes is difficult, since the tradeoff changes based on node mobility rate and communication patterns. This has led to the development of a new family of routing protocols that combine proactive and a reactive routing algorithms. These new protocols automatically adjust the radius of proactive information dissemination to discover routes with low overhead and latency.

3 DoD Interactions and Technology Transitions

A variety of technology transitions and interactions with DoD organization occurred during the period of this funding:

- Schneider chaired a study for DARPA IPTO Program Manager Jay Lala on promising research directions for Self-Healing Networked Information Systems.
- Schneider chaired the DARPA IPTO Oasis Dem-Val External Evaluation Committee.
- Morrisett and Schneider worked with Microsoft to develop a .NET in-lined reference monitor (IRM).
- Researchers at Carnegie-Mellon University, Princeton University, University of California (Riverside), University of Newcastle-Upon-Tyne, and Intel Research are all now building on PoET/PSLang IRM tools developed by Schneider and collaborators.
- Further public releases of the Jif compiler have been made available at the Jif web site, http://www.cs.cornell.edu/jif. The Jif language extends the Java programming language with support for information flow control. The Jif compiler is implemented on top of the Polyglot extensible compiler framework for Java. The Polyglot framework has also been released publicly at http://www.cs.cornell.edu/projects/polyglot, and researchers at Princeton University are using this framework in their own research. The releases of both Jif and Polyglot are provided as Java source code and work on Unix and Windows platforms.
- AT&T research is collaborating to develop the Cyclone language, compiler, and tools. In addition, researchers at the University of Maryland, the University of Utah, Princeton, and the University of Pennsylvania, and Cornell are all using Cyclone to develop research prototypes.

4 Publications Supported under this Grant

(1) Stuart Allen, Robert L. Constable, Richard Eaton, Christoph Kreitz, and Lori Lorigo. The Nuprl Open Logical Environment. *17th International Conference on Automated Deduction*, volume 1831 of *Lecture Notes in Artificial Intelligence*, Springer Verlag, (2000).

(2) B. Atkin and K. Birman. Evaluation of an Adaptive Transport Protocol. *INFOCOM 2003*. Submitted, July 2002.

(3) B. Atkin, E. G. Sirer. PortOS: An Educational Operating System for the Post-PC Environment. *33rd ACM Technical Symposium on Computer Science Education (SIGCSE)* (Cincinnati, Ohio, February 2002).

(4) J. Ayres, J. E. Gehrke, T. Yiu and J. Flannick. Sequential Pattern Mining Using Bitmaps. *In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Edmonton, Alberta, Canada, July 2002).

(5) R. Barr, J.C. Bicket, D.S. Dantas, B. Du, T.W. Danny Kim, B. Zhou and E. Gun Sirer. On the Need for System-Level Support for Ad hoc and Sensor Networks. *Operating Systems Review 36*, 2 (April 2002), ACM, 1–5.

(6) S. Ben-David, J. E. Gehrke and R. Schuller. A Theoretical Framework for Learning from a Pool of Disparate Data Sources. *In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Edmonton, Alberta, Canada, July 2002).

(7) Mark Bickford, Christoph Kreitz, Robbert van Renesse, and Robert Constable. An Experiment in Formal Design Using Meta-properties. *Proceedings of DARPA Information Survivability Conference & Exposition II (DISCEX 2001)*, IEEE Computer Society Press, June 2001.

(8) Mark Bickford, Christoph Kreitz, Robbert van Renesse, and Xiaoming Liu. Proving Hybrid Protocols Correct. *Proceedings of the 14th International Conference on Theorem Proving in Higher Order Logics*, Lecture Notes in Computer Science, Volume 2152 (G. Goos, J. Hartmanis, J. van Leeuwen, editors), Springer-Verlag, September 2001.

(9) Kenneth P. Birman. The Next Generation Internet: Unsafe at Any Speed? *IEEE Computer, Special Issue on Survivability of the Nationally Critical Infrastructure*, Anita Jones, Ed. (August 2000).

(10) Kenneth P. Birman. Technology challenges for virtual overlay networks. *IEEE Transactions on Systems, Man and Cybernetics 31*, 4 (July 2001), 319–327.

(11) Kenneth P. Birman and Oznur Ozkasap. Throughput Stability of Reliable Multicast Protocols. *ADVIS' 2000*, (Dokuz Eylul University, Izmir, Turkey, October 2000).

(12) K. Birman and R. van Renesse. Scalable Data Fusion Using Astrolabe. *Fifth International Conference on Information Fusion 2002 (IF 2002).* Submitted, January 2001.

(13) Kenneth P. Birman, Robbert van Renesse, and Werner Vogels. Spinglass: Secure and Scalable Communication Tools For Mission-Critical Computing. *Proceedings of DARPA Information Survivability Conference & Exposition II (DISCEX 2001)*, IEEE Computer Society Press, June 2001, 85–99.

(14) Kenneth P. Birman, Werner Vogels, and Robbert van Renesse. Using Epidemic Techniques for Building Ultra-Scalable Reliable Communications Systems. *Workshop on New Visions for Large-Scale Networks: Research and Applications* (Vienna, Virginia, March 2001).

(15) Philippe Bonnet, Alan Demers, and Praveen Seshadri. Towards Sensor Database Systems. *Proceedings of the Second International Conference on Mobile Data Management*, (Hong Kong, China, January 2001).

(16) Philippe Bonnet, Johannes Gehrke, and Praveen Seshadri. Querying the Physical World. *IEEE Personal Communications, Special Issue on Smart Spaces and Environments*, (October 2000).

(17) P.S. Bradley, J. E. Gehrke, R. Ramakrishnan and R. Srikant. Philosophies and Advances in Scaling Mining Algorithms to Large Databases. *Communications of the ACM*, August 2002.

(18) C. Bucila, J.E. Gehrke, D. Kifer and W. White. DualMiner: A Dual-Pruning Algorithm for Itemsets with Constraints. *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Edmonton, Alberta, Canada, July 2002).

(19) Doug Burdick, Manuel Calimlim, and Johannes Gehrke. MAFIA: A Maximal Frequent Itemset Algorithm for Transactional Databases. *Proceedings of the 17th International Conference on Data Engineering* (Heidelberg, Germany, April 2001), 443–452.

(20) Ranveer Chandra, Vanugupalen Ramasubramanian, and Kenneth P. Birman. Anonymous Gossip: Improving Multicast Reliability in Ad-Hoc Networks. *Proceedings International Conference on Distributed Computing Systems (ICDCS 2001)* (Phoenix, Arizona April 2001), 275–283.

(21) Y. Chen, E. G. Sirer and S. Wicker. On The Determination of Optimal Transmission Radius in Ad Hoc Networks. *Hawaii International Conference on Systems Sciences*. To appear, January 2003.

(22) Z. Chen. *Building Compressed Database Systems*. Ph.D. Thesis, Cornell University, August 2002.

(23) J. Cheney and R. Hinze. Poor man's generics and dynamics. To appear, *Haskell Workshop 2002*.

(24) Francis Chu and Joseph Y. Halpern. A Decision-theoretic Approach to Reliable Message Delivery. *Distributed Computing, 14*, 1 (2001), 1–16.

(25) F. Chu, J. Halpern and J. E. Gehrke. Least Expected Cost Query Optimization: What Can We Expect? *Proceedings of the 21st ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS 2002)* (Madison, Wisconsin, June 2002).

(26) R. L. Constable. Naive Computational Type Theory. *Proc. International Summer School on Proof and System-Reliability*, NATO Science Series III, Kluwer (2002).

(27) R. L. Constable. Writing constructive Proofs Yielding Efficient Extracted Programs. *Electronic Notes in Theoretical Computer Science 37* (2001).

(28) R. Constable and K. Crary. Computational Complexity and Induction for Partial Computable Functions in Type Theory. *Association for Symbolic Logic*, Essays in Honor of Solomon Feferman, (2001).

(29) Robert L Constable and Jason Hickey. Nuprl's Class Theory and its Applications. *Foundations of Secure Computation* (F.L. Bauer and R. Steinbruggen, eds.), IOS Press, 2000, 91–115.

(30) Robert L. Constable, P.B. Jackson, P. Naumov, and J. Uribe. Constructively Formalizing Automata. *Proof, Language and Interaction: Essays in Honour of Robin Milner*, MIT Press, 2000, 213-238.

(31) K. Crary, S. Weirich and G. Morrisett. Intensional polymorphism in type erasure semantics. *Journal of Functional Programming.* To appear.

(32) D. V. Coury, J. S. Thorp, K. M. Hopkinson and K. P. Birman. An Agent-based Current Differential Relay for use with a Utility Intranet. *IEEE Transactions on Power Delivery 17*, 1 (January 22, 2002), 47–53.

(33) Denis V. Coury, Jim S. Thorp, Kenneth M. Hopkinson, and Kenneth P. Birman. Agent Technology Applied to Adaptive Relay Settings for Multi-Terminal Lines. *IEEE Power Engineering Society Summer Meeting 2* (2000), 1196–1201.

(34) A. Das, I. Gupta and A. Motivala. SWIM: Scalable Weakly-consistent Infection-style Process Group Membership Protocol. *Proceedings of the International Conference on Dependable Systems and Networks 2002 (DSN 2002)*, June 2002, 303-312.

(35) Alin Dobra and Johannes Gehrke. Bias Correction in Classification Tree Construction. *Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2001)*, (Williams College, Mass, June 2001).

(36) A. Dobra, M. Garofalakis, J. E. Gehrke and R. Rastogi. Processing Complex Aggregate Queries over Data Streams. *Proceedings of the 2002 ACM Sigmod International Conference on Management of Data (SIGMOD 2002)* (Madison, Wisconsin, June 2002).

(37) A. Dobra and J. E. Gehrke. SECRET: A Scalable Linear Regression Tree Algorithm. *In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Edmonton, Alberta, Canada, July 2002).

(38) U. Erlingsson and F. B. Schneider. IRM Enforcement of Java Stack Inspection. *Proceedings 2000 IEEE Symposium on Security and Privacy* (Oakland, CA, May 2000), IEEE Computer Society, Los Alamitos, CA, 246–255.

(39) A. Evfimievski, R. Srikant, R. Agrawal and J. Gehrke. Privacy Preserving Mining of Association Rules. *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Edmonton, Alberta, Canada, July 2002).

(40) S.A. Fakhouri, G. Goldszmidt, I. Gupta, M. Kalantar and J. Pershing. GulfStream – a System for Dynamic Topology Management in Multi-Domain Server Farms. *Proc. IEEE International Conference on Cluster Computing (Cluster 2001)*, October 2001.

(41) A. Faradjian, J. E. Gehrke and P. Bonnet. GADT: A Probability Space ADT For Representing and Querying the Physical World. *Proceedings of the 18th International Conference on Data Engineering (ICDE 2002)* (San Jose, California, February 2002).

(42) K. Fisher, R. Pucella and J. Reppy. A framework for interoperability. *Proceedings of the First International Workshop on Multi-Language Infrastructure and Interoperability (BABEL'01)*, Electronic Notes in Theoretical Computer Science, 59(1), 2001.

(43) Nir Friedman, Joseph Y. Halpern, and D. Koller. First-order Conditional Logic Revisited. *ACM Transactions on Computational Logic 1*, 2 (2000), 175–207.

(44) W. F. Fung, D. Sun and J. E. Gehrke. COUGAR: The Network is the Database. *Proceedings of the 2002 ACM Sigmod International Conference on Management of Data (SIGMOD 2002)* (Madison, Wisconsin, June 2002). Demo description.

(45) V. Ganti, J. E. Gehrke, and Raghu Ramakrishnan. DEMON: Mining and Monitoring Evolving Data. *IEEE Transactions on Knowledge and Data Engineering 13*, 1 (January/February 2001), 50–63.

(46) V. Ganti, J. E. Gehrke and R. Ramakrishnan. Mining Data Streams under Block Evolution. *SIGKDD Explorations 3*, 2 (January 2002). Invited paper.

(47) V. Ganti, J. E. Gehrke, R. Ramakrishnan and W.-Y. Loh. A Framework for Measuring Changes in Data Characteristics. *Journal of Computer and System Sciences 64*, 3 (May 2002), 542–578.

(48) J. E. Gehrke. Report on the SIGKDD 2001 Conference Panel New Research Directions in KDD. *SIGKDD Explorations 3*, 2 (January 2002).

(49) Johannes Gehrke, Flip Korn, and Divesh Srivastava. On Computing Correlated Aggregates Over Continual Data Streams. *Proceedings of the 2001 ACM SIGMOD International Conference on Management of Data* (Santa Barbara, California, May 2001).

(50) Johannes Gehrke, Raghu Ramakrishnan, Venkatesh Ganti. RAINFOREST – A Framework for Fast Decision Tree Construction of Large Datasets. *Data Mining and Knowledge Discovery 4*, 2/3 (July 2000), 127–162.

(51) David Gries and Fred B. Schneider. Formalizations of substitutions of equals for equals. *Millennial Perspectives in Computer Science, Proceedings of the 1999 Oxford-Microsoft Symposium in honour of Professor Sir Anthony Hoare*, (Davies, Roscoe, and Woodcock, editors), Palgrave Publishers, Hampshire, England, November 2000, 119–132.

(52) D. Grossman. Existential Types for Imperative Languages. Type checking systems code. *Eleventh European Symposium on Programming* (Grenoble, France, April 2002), Lecture Notes in Computer Science, Volume 2305, 21–35.

(53) Dan Grossman and Greg Morrisett. Scalable Certification for Typed Assembly Language. *2000 ACM SIGPLAN Workshop on Types in Compilation*, Lecture Notes in Computer Science, Volume 2071 (Robert Harper, editor), Springer-Verlag, Montreal, 2000, 117–146.

(54) Dan Grossman, Steve Zdancewic, and Greg Morrisett. Syntactic Type Abstraction. *ACM Transactions on Programming Languages and Systems 22*, 6 (November 2000), 1037-1080.

(55) D. Grossman, G. Morrisett, T. Jim, M. Hicks, J. Cheney, and Y. Wang. Region-based memory management in Cyclone. *ACM Conference on Programming Language Design and Implementation* (Berlin, Germany, June 2002), 282–293.

(56) P. Grunwald and J. Y. Halpern. Updating probabilities. *Proceedings of the Eighteenth Conference on Uncertainty in AI* (2002), 187–196.

(57) I. Gupta, K. P. Birman and R. van Renesse. Fighting Fire with Fire: Using Randomized Gossip to Combat Stochastic Scalability Limits. *Journal of Quality and Reliability Engineering International 18*, 3 (May/June 2002), Special Issue on Quality and Reliability of Computer Network Systems. Nong Ye (editor), 165–184.

(58) I. Gupta, A. Kermarrec and A.J. Ganesh. Efficient Epidemic-Style Protocols for Reliable and Scalable Multicast. *Proceedings of the International Workshop on Reliable Peer-to-Peer Systems* (Osaka, Japan, October 2002).

(59) Indranil Gupta, Robbert van Renesse, and Kenneth P. Birman. Scalable fault-tolerant aggregation in large process groups. *Proceedings of International Conference on Dependable Systems and Networks*, (Goteborg, Sweden, July 2001), 433–442.

(60) Indranil Gupta, Robbert van Renesse, and Kenneth P. Birman. A Probabilistically Correct Election Protocol for Large Groups. *DISC 2000* (Toledo, Spain, October 2000).

(61) Z. Haas, J. Y. Halpern and L. Li. Gossip-based ad hoc routing. *Proceedings of Infocom* 2002, 1707–1716.

(62) Z. Haas, L. Li, Joseph Y. Halpern, and S.B. Wicker. A Decision-theoretic Approach to Resource Allocation in Wireless Multimedia Networks. *Proceedings Dial M for Mobility* (2001), 86–95.

(63) Joseph Y. Halpern. A Note on Knowledge-based Programs and Specifications. *Distributed Computing 13*, 3 (2000), 145–153.

(64) Joseph Y. Halpern. Conditional Plausibility Measures and Bayesian Networks. *Proceedings of the Sixteenth Conference on Uncertainty in AI* (2000) 247–255.

(65) Joseph Y. Halpern. Axiomatizing Causal Reasoning. *Journal of AI Research 12* (2000), 317–337.

(66) Joseph Y. Halpern. Review of 'Probability and Conditionals: Belief Revision and Rational Decisions.' *Journal of Philosophical Logic 100*, 2 (2000), 277–281.

(67) Joseph Y. Halpern. Editorial: An Author's Bill of Rights and Responsibilities. *Journal of the ACM 47*, 5 (2000), 828–825.

(68) Joseph Y. Halpern, Robert Harper, Neil Immerman, Phokion G. Kolaitis, Moshe Y. Vardi, and Victor Vianu. The Unusual Effectiveness of Logic in Computer Science. *Bulletin of Symbolic Logic 7*, 2 (2001), 213–236.

(69) Joseph Y. Halpern and G. Lakemeyer. Multi-agent Only Knowing. *Journal of Logic and Computation 11*, 1 (2001), 41–70.

(70) J. Y. Halpern, Z. Haas, L. Li, and S. B. Wicker. A Decision Theoretic-approach to Resource Allocation in Wireless Multimedia Networks. *Proceedings of Dial M for Mobility* (2000), 86–95.

(71) J. Y. Halpern and K. O'Neill. Secrecy in multi-agent systems. *Proceedings of the 15th IEEE Computer Security Foundations Workshop* 2002, 32–46.

(72) J. Y. Halpern and R. Pucella. A logic for reasoning about upper probabilities. *Journal of AI Research 17* (2002), 57–81.

(73) J. Y. Halpern and R. Pucella. Reasoning about Expectation. *Proceedings of the Eighteenth Conference on Uncertainty in AI* (2002), 207–215.

(74) J. Y. Halpern and R. Pucella. On the relationship between strand spaces and multi-agent systems. *ACM Transactions on Information and System Security*. To appear.

(75) Joseph Y. Halpern and R. van der Meyden. A Logical Reconstruction of SPKI. *Proceedings of the 14th IEEE Computer Security Foundations Workshop* (Cape Breton, Nova Scotia, Canada, June 2001).

(76) David Harel, Dexter Kozen, and Jerzy Tiuryn. *Dynamic Logic*, MIT Press, Cambridge, MA (2000).

(77) D. Harel, D. Kozen, and J. Tiuryn. Dynamic logic. (D. M. Gabbay and F. Guenthner, editors) *Handbook of Philosophical Logic*, Volume 4, Kluwer, 2nd edition, 2002, 99–217.

(78) Kate Jenkins and Alan Demers. Logarithmic Harary Graphs. *International Workshop on Applied Reliable Group Communication (WARGC 2001)* (Phoenix, Arizona, April 2001).

(79) Kate Jenkins, K. Hopkinson, Kenneth P. Birman. A Gossip Protocol for Subgroup Multicast. *International Workshop on Applied Reliable Group Communication (WARGC 2001)* (Phoenix, Arizona, April 2001).

(80) T. Jim, G. Morrisett, D. Grossman, M. Hicks, J. Cheney, and Y. Wang. Cyclone: A safe dialect of C. *Usenix Annual Technical Conference* (Monterey, CA, June 2002).

(81) D. Kempe, J. Kleinberg, and A. Demers. Spatial Gossip and Resource Location Protocols. *Proceedings of the 33rd ACM Symposium on Theory of Computing* (Crete, Greece, July 2001).

(82) D. Kozen. Certification of Compiler Optimizations using Kleene Algebra with Tests. *Proc. 1st Int. Conf. Computational Logic (CL2000)*, Vol. 1861 of *Lecture Notes in Artificial Intelligence*, J. Lloyd, V. Dahl, U. Furbach, M. Kerber, K. -K. Lau, C. Palamidessi, L. M. Pereira, Y. Sagiv, and P. J. Stuckey, Eds., London, Springer-Verlag, (July 2000), 568-582.

(83) D. Kozen. Computational inductive definability. Submitted for publication.

(84) D. Kozen. On the complexity of reasoning in Kleene algebra. *Information and Computation*. To appear.

(85) D. Kozen. On two letters versus three. In Zoltán Ésik and Anna Ingólfsdóttir, editors, *Proceedings of Workshop on Fixed Points in Computer Science* (FICS'02), July 2002, 44–50.

(86) D. Kozen. Some results in dynamic model theory (abstract). In E. A. Boiten and B. Möller, editors, *Proceedings Conference on Mathematics of Program Construction* (MPC'02), Lecture Notes in Computer Science, Volume 2386 (July 2002), 21.

(87) Dexter Kozen. Myhill-Nerode Relations on Automatic Systems and the Completeness of Kleene Algebra. *Proc. 18th Symp. Theoretical Aspects of Computer Science*, (Dresden, Germany, February 2001), Lecture Notes in Computer Science, volume 2010 (A. Ferreira and H. Reichel, editors), Springer-Verlag, Heidelberg, 27–38.

(88) D. Kozen and M. Stillerman. Eager class initialization for Java. *Proceedings 7th International Symposium on Formal Techniques in Real-Time and Fault Tolerant Systems (FTRTFT'02) IFIP* (Germany, Sept. 2002). To appear.

(89) D. Kozen and Jerzy Tiuryn. On the completeness of propositional Hoare logic. *Information Sciences 139*, 2001, 187–195.

(90) Dexter Kozen and Jerzy Tiuryn. Intuitionistic Linear Logic and Partial Correctness. *Proc. 16th Symp. Logic in Computer Science* (June 2001), IEEE Press.

(91) C. Kreitz. Building Reliable, High-Performance Networks with the Nuprl Proof Development System. *Journal of Functional Programming* Submitted for publication.

(92) C. Kreitz. Designing Reliable, High-Performance Networks with the Nuprl Logical Programming Environment. *2002 AAAI Spring Symposium on Logic-Based Program Synthesis*, March 2002.

(93) C. Kreitz and B. Pientka. Connection-based Inductive Theorem Proving. *Studia Logica (2001)*.

(94) C. Kreitz and H. Mantel. A Matrix Characterization for Multiplicative Exponential Linear Logic. *Journal of Automated Reasoning*. Submitted for publication, July 2002.

(95) Christoph Kreitz, Jens Otten, Stephan Schmitt, and Brigitte Pientka, Matrix-based Constructive Theorem Proving. *Intellectics and Computational Logic: Papers in honor of Wolfgang Bibel*, Applied Logic Series 19, Kluwer, April 2000.

(96) Christoph Kreitz and Brigitte Pientka. Matrix-based Inductive Theorem Proving. *International Conference TABLEAUX-2000*, volume 1847 of *Lecture Notes in Artificial Intelligence*, R. Dyckhoff, Ed.Springer Verlag, 2000, 294–308.

(97) Christoph Kreitz and Brigitte Pientka. Connection-based Inductive Theorem Proving. *Studia Logica*, 2001.

(98) C. Kreitz and S. Schmitt. A Uniform Procedure for Converting Matrix Proofs into Sequent-Style Systems. *Journal of Information and Computation 162*, 1–2 (2000), 226–254.

(99) Li Li and Joseph Y. Halpern. Minimum-energy Mobile Wireless Networks Revisited. *Proceedings of the IEEE Conference on Communications*, 1 (2001), 278–283.

(100) Xiaoming Liu and Robbert van Renesse. Fast Protocol Transition in A Distributed Environment. *Proc. of 19th ACM Conference on Principles of Distributed Computing (PODC 2000)*, (Portland, OR, July 2000).

(101) X. Liu, R. van Renesse, M. Bickford, C. Kreitz, and Robert Constable. Protocol Switching: Exploiting Meta-properties. *Proceedings of the International Workshop on Applied Reliable Group Communication (WARGC 2001)*, (Phoenix, Arizona, April 2001), IEEE Computer Society Press.

(102) G. McGraw and G. Morrisett. Attacking Malicious Code: A Report to the Infosec Research Council. *IEEE Software 17*, 5 (Sept.-Oct. 2000), 33–41.

(103) Heiko Mantel and Christoph Kreitz. A Matrix Characterization for Multiplicative Exponential Linear Logic. *Journal of Theoretical Computer Science*. Submitted for publication, April 2000.

(104) Lynette I. Millett and Tim Teitelbaum. Issues in Slicing Promela and its Applications to Model Checking, Protocol Understanding, and Simulation. *International Journal on Software Tools for Technology Transfer 2*, 4 (2000), 343–349.

(105) G. Morrisett. Type checking systems code. *Eleventh European Symposium on Programming* (Grenoble, France, April 2002), Lecture Notes in Computer Science, Volume 2305, 1–5.

(106) G. Morrisett, K. Crary, N. Glew, and D. Walker. Stack-based typed assembly language. *Journal of Functional Programming 12*, 1 (January 2002), University Press, Cambridge, England, 43–88.

(107) Andrew Myers. Protecting Privacy using the Decentralized Label Model. *ACM Transactions on Software Engineering Methodology 9*, 4 (2000), 410–442.

(108) Wei Tsang Ooi, Robbert van Renesse, Brian Smith. Design and Implementation of Programmable Media Gateways. *10th International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV 2000)*, (Chapel Hill, NC, June 2000).

(109) P. Papadimitratos, Z. Haas, E. G. Sirer. Path-Set Selection in Mobile Ad Hoc Networks. *Proceedings of the Third ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc)* (Lausanne, Switzerland, June 2002).

(110) R. Pucella and V. Weissman. A logic for reasoning about digital rights. *Proceedings of the 15th IEEE Computer Security Foundations Workshop* 2002, 282–294.

(111) R. Ramakrishnan and J. Gehrke. *Database Management Systems*. McGraw Hill Higher Education, Third Edition, 2002.

(112) V. Ramasubramanian and D. Mosse. SRL: Providing a Birdirectional Abstraction for Unidirectional Ad Hoc Networks. *Transactions on Networking*. Submitted for publication, 2002.

(113) V. Ramasubramanian and D. Mosse. Statistical Analysis of Connectivity in Unidirectional Mobile Ad Hoc Networks. *International Workshop on Ad Hoc Networking 2002* (Vancouver, Canada, August 18-21, 2002).

(114) V. Ramasubramanian, R. Chandra and D. Mosse. Providing a Bidirectional Abstraction for Unidirectional Ad Hoc Networks. *INFOCOM 2002* (New York, NY, June 23-27, 2002).

(115) O. Rodeh and K. Birman. A Simulation Model for an Epidemic Multicast Protocol. *Proceedings of the BAS2000 Conference (5th Computer Networks Symposium)* (Bilkent University, Ankara, Turkey, June 2000).

(116) O. Rodeh, K. Birman and D. Dolev. The Architecture and Performance of the Security Protocols in the Ensemble Group Communication System. *ACM Transactions on Information Systems and Security (TISSEC) 4*, 3 (August 2001), 289–319.

(117) O. Rodeh, K. P. Birman and D. Dolev. Using AVL Trees for Fault-Tolerant Group Key Management. *International Journal of Information Security (IJIS) 1*, 2 (February 2002), 84–99.

(118) Lus Rodrigues, Katherine Guo, Paulo Verssimo, Kenneth P. Birman. A Dynamic Light-Weight Group Service. *Journal of Parallel and Distributed Computing 60* (Dec. 2000), 1449.

(119) A. Sabelfeld and A. C. Myers. End-to-end security via program analysis. Submitted for publication.

(120) Stephan Schmitt, Lori Lorigo, Christoph Kreitz, and Alexey Nogin. JProver: Integrating Connection-based Theorem Proving into Interactive Proof Assistants. *Proceedings of the International Joint Conference on Automated Reasoning*, Lecture Notes in Artificial Intelligence, Volume 2083, (J. Carbonell and J. Siekmann, editors), Springer-Verlag, June 2001, 421.

(121) F. B. Schneider. Open Source in Security: Visiting the Bizarre. *Proceedings 2000 IEEE Symposium on Security and Privacy* (Oakland, CA, May 2000), IEEE Computer Society, Los Alamitos, CA, 126–127.

(122) Fred B. Schneider Editorial: Time for Change. *Distributed Computing 13*, 4 (November 2000), 187.

(123) Fred B. Schneider, Greg Morrisett, and Robert Harper. A language-based approach to security. *Informatics: 10 Years Back, 10 Years Ahead*. Lecture Notes in Computer Science, Volume 2000 Reinhard Wilhelm, editor, Springer Verlag, Heidelberg, 2000, 86–101.

(124) J. Shanmugasundaram, J. Kiernan, E. Shekita, C. Fan, J. Funderburk. Querying XML Views of Relational Data. *Proceedings of the International Conference on Very Large Data Bases (VLDB)* (Rome, Italy, September 2001).

(125) J. Shanmugasundaram, E. Shekita, J. Kiernan, R. Krishnamurthy, E. Viglas, J. Naughton, I. Tatarinov. A General Technique for Querying XML Documents Using a Relational Database Systems. *SIGMOD Record, 30*, 2 (September 2001).

(126) E. G. Sirer, K. Wang. A Temporal Logic-Based Access Control Mechanism for Web Services. *Proceedings of the Eighth Symposium on Access Control Models and Technologies (SACMAT)* (Monterey, California, May 2002).

(127) F. Smith. *Certified Run-Time Code Generation*. Ph.D. Thesis, Cornell University, January 2002.

(128) Frederick Smith, David Walker, and Greg Morrisett. Alias Types. *European Symposium on Programming* (Berlin, Germany, March 2000).

(129) F. Smith, D. Grossman, G. Morrisett, L. Hornof, and T. Jim. Compiling for template-based run-time code generation. To appear, *Journal of Functional Programming*.

(130) I. Tatarinov, E. Viglas, K. Beyer, J. Shanmugasundaram, E. Shekita, C. Zhang. Storing and Querying Ordered XML Using a Relational Database System. *Proceedings of the ACM SIGMOD International Conference on Management of Data* (Madison, Wisconsin, May 2002).

(131) R. van Renesse. The Importance of Aggregation. *In Proceedings of the International Workshop on Future Directions in Distributed Computing* (Bertinoro, Italy, June 2002).

(132) R. van Renesse. Power-Aware Epidemics. *In Proceedings of the International Workshop on Reliable Peer-to-Peer Systems* (Osaka, Japan, October 2002).

(133) Robbert van Renesse, Kenneth P. Birman, and Werner Vogels. Using Epidemic Techniques for Building Ultra-scalable Reliable Communication Systems. *Workshop on New Visions for Large-Scale Networks: Research and Applications* (Vienna, Virginia, March 2001).

(134) Robbert van Renesse, Kenneth P. Birman, and Werner Vogels. Spinglass, Scalable and Secure Communication Tools for Mission-critical Computing. *Proceedings of DARPA Information Survivability Conference & Exposition II (DISCEX 2001)*, IEEE Computer Society Press, June 2001.

(135) R. van Renesse, D. Dumitriu. Collaborative Networking in an Uncooperative Internet. *In Proceedings of the 21st Symposium on Reliable Distributed Systems* (Osaka, Japan, October 2002).

(136) R. van Renesse, K. P. Birman, D. Dumitriu and W. Vogels. Scalable Management and Data Mining Using Astrolabe. *Proceedings of the First International Workshop on Peer-to-Peer Systems (IPTPS)* (Cambridge, Massachusetts, March 2002).

(137) R. van Renesse, and K. Birman. Astrolabe: A Robust and Scalable Technology for Distributed System Monitoring, Management, and Data Mining. Submitted to *ACM TOCS*.

(138) W. Vogels, C. Re, R. van Renesse and K. Birman. A Collaborative Infrastructure for Scalable and Robust News Delivery. *Proceedings of the IEEE Workshop on Resource Sharing in Massively Distributed Systems (RESH'02)* (Vienna, Austria, July 2002).

(139) W. Vogels, R. van Renesse and K. Birman. Collaborative Content Delivery: A Peer-to-Peer Solution for Web-Based Publish/Subscribe. *First International Workshop on Peer-to-Peer Systems(IPTPS 2002)* (Cambridge, Massachusetts, March 2002).

(140) David Walker and Greg Morrisett. Alias Types for Recursive Data Structures. *ACM SIGPLAN Workshop on Types in Compilation*, Lecture Notes in Computer Science, Volume 2071, (Robert Harper, editor), Springer-Verlag, Montreal, 2000, 177–206.

(141) X. Wang, K. Hopkinson, J. Thorp, R. Giovanini, K. Birman and D. Coury. Developing an Agent-based Backup Protection System for Transmission Networks. *Power Systems and Communications Infrastructures for the Future Conference* (Beijing, China). Submitted, September 2002.

(142) S. Weirich. Higher-order intensional type analysis. *Eleventh European Symposium on Programming* (Grenoble, France, April 2002), Lecture Notes in Computer Science, Volume 2305, 98–114.

(143) S. Weirich. *Programming With Types*. Ph.D. Thesis, Cornell University, July 2002.

(144) Zhen Xiao and Kenneth P. Birman. Providing Efficient, Robust Error Recovery Through Randomization. *International Workshop on Applied Reliable Group Communication (WARGC 2001)* (Phoenix, Arizona, April 2001), 31–36.

(145) Zhen Xiao and Kenneth P. Birman. A Randomized Error Recovery Algorithm for Reliable Multicast. *IEEE Infocom 2001* (Alaska, April 2001).

(146) Z. Xiao, R. van Renesse and K. Birman. Optimizing Buffer Management for Reliable Multicast. *Dependable Systems and Networks (DSN '02)* (Bethesda, Maryland, July 2002), IEEE.

(147) S. Zdancewic and A. C. Myers. Secure information flow and linear continuations. *Higher Order and Symbolic Computation, 15* (2-3), 2002.

(148) Steve Zdancewic and Andrew C. Myers. Robust Declassification. *Proceedings of the 14th IEEE Computer Security Foundations Workshop* (Cape Breton, Nova Scotia, Canada, June 2001).

(149) Steve Zdancewic and Andrew C. Myers. Secure Information Flow and CPS. *Proceedings of the 10th European Symposium on Programming* (Genova, Italy, April 2001).

(150) S. Zdancewic, L. Zheng, N. Nystrom, and A. C. Myers. Untrusted hosts and confidentiality: Secure program partitioning. *Proceedings of the 18th ACM Symposium on Operating Systems Principles* (Banff, Canada, October 2001), 1–14.

(151) S. Zdancewic, L. Zheng, N. Nystrom, and A. C. Myers. Secure program partitioning. *ACM Transactions on Computing Systems 20*, 3 (August 2002), 282–328.

(152) L. Zheng, S. Chong, S. Zdancewic, and A. C. Myers. Enforcing end-to-end integrity with replicated code partitions. Submitted for publication, May 2002.

(153) L. Zhou, R. van Renesse and M. Marsh. Implementing IPv6 as a Peer-to-Peer Overlay Network. *In Proceedings of the International Workshop on Reliable Peer-to-Peer Systems* (Osaka, Japan, October 2002).